

-1-

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Human Genome Sciences, Inc.
- (ii) TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
- (iii) NUMBER OF SEQUENCES: 35
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
  - (B) STREET: 9410 KEY WEST AVENUE
  - (C) CITY: ROCKVILLE
  - (D) STATE: MARYLAND
  - (E) COUNTRY: USA
  - (F) ZIP: 20850
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: TO BE ASSIGNED
  - (B) FILING DATE: HEREWITH
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 09/042,105
  - (B) FILING DATE: 13-MAR-1998
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 09/107,997
  - (B) FILING DATE: 30-JUN-1998
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: MICHELE M. WALES
  - (B) REGISTRATION NUMBER: 43,975
  - (C) REFERENCE/DOCKET NUMBER: PF112PCT3
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (301) 309-8504
  - (B) TELEFAX: (301) 309-8439

-2-

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: sig\_peptide
- (B) LOCATION: 12..80

## (ix) FEATURE:

- (A) NAME/KEY: mat\_peptide
- (B) LOCATION: 81..1268

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 12..1268

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

GTCCTTCCAC C ATG CAC TCG CTG GGC TTC TTC TCT GTG GCG TGT TCT CTG      50
      Met His Ser Leu Gly Phe Phe Ser Val Ala Cys Ser Leu
      -23          -20          -15

CTC GCC GCT GCG CTG CTC CCG GGT CCT CGC GAG GCG CCC GCC GCC GCC      98
Leu Ala Ala Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala
-10          -5          1          5

GCC GCC TTC GAG TCC GGA CTC GAC CTC TCG GAC GCG GAG CCC GAC GCG      146
Ala Ala Phe Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala
      10          15          20

GGC GAG GCC ACG GCT TAT GCA AGC AAA GAT CTG GAG GAG CAG TTA CGG      194
Gly Glu Ala Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg
      25          30          35

TCT GTG TCC AGT GTA GAT GAA CTC ATG ACT GTA CTC TAC CCA GAA TAT      242
Ser Val Ser Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr
      40          45          50

TGG AAA ATG TAC AAG TGT CAG CTA AGG AAA GGA GGC TGG CAA CAT AAC      290
Trp Lys Met Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn
      55          60          65          70

AGA GAA CAG GCC AAC CTC AAC TCA AGG ACA GAA GAG ACT ATA AAA TTT      338
Arg Glu Gln Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe
      75          80          85

```

-3-

GCT GCA GCA CAT TAT AAT ACA GAG ATC TTG AAA AGT ATT GAT AAT GAG Ala Ala Ala His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu	386
90 95 100	
TGG AGA AAG ACT CAA TGC ATG CCA CGG GAG GTG TGT ATA GAT GTG GGG Trp Arg Lys Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly	434
105 110 115	
AAG GAG TTT GGA GTC GCG ACA AAC ACC TTC TTT AAA CCT CCA TGT GTG Lys Glu Phe Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val	482
120 125 130	
TCC GTC TAC AGA TGT GGG GGT TGC TGC AAT AGT GAG GGG CTG CAG TGC Ser Val Tyr Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys	530
135 140 145 150	
ATG AAC ACC AGC ACG AGC TAC CTC AGC AAG ACG TTA TTT GAA ATT ACA Met Asn Thr Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr	578
155 160 165	
GTG CCT CTC TCT CAA GGC CCC AAA CCA GTA ACA ATC AGT TTT GCC AAT Val Pro Leu Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn	626
170 175 180	
CAC ACT TCC TGC CGA TGC ATG TCT AAA CTG GAT GTT TAC AGA CAA GTT His Thr Ser Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val	674
185 190 195	
CAT TCC ATT ATT AGA CGT TCC CTG CCA GCA ACA CTA CCA CAG TGT CAG His Ser Ile Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln	722
200 205 210	
GCA GCG AAC AAG ACC TGC CCC ACC AAT TAC ATG TGG AAT AAT CAC ATC Ala Ala Asn Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile	770
215 220 225 230	
TGC AGA TGC CTG GCT CAG GAA GAT TTT ATG TTT TCC TCG GAT GCT GGA Cys Arg Cys Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly	818
235 240 245	
GAT GAC TCA ACA GAT GGA TTC CAT GAC ATC TGT GGA CCA AAC AAG GAG Asp Asp Ser Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu	866
250 255 260	
CTG GAT GAA GAG ACC TGT CAG TGT GTC TGC AGA GCG GGG CTT CGG CCT Leu Asp Glu Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro	914
265 270 275	
GCC AGC TGT GGA CCC CAC AAA GAA CTA GAC AGA AAC TCA TGC CAG TGT Ala Ser Cys Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys	962
280 285 290	
GTC TGT AAA AAC AAA CTC TTC CCC AGC CAA TGT GGG GCC AAC CGA GAA Val Cys Lys Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu	1010
295 300 305 310	

-4-

TTT GAT GAA AAC ACA TGC CAG TGT GTA TGT AAA AGA ACC TGC CCC AGA	1058
Phe Asp Glu Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg	
315 320 325	
AAT CAA CCC CTA AAT CCT GGA AAA TGT GCC TGT GAA TGT ACA GAA AGT	1106
Asn Gln Pro Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser	
330 335 340	
CCA CAG AAA TGC TTG TTA AAA GGA AAG AAG TTC CAC CAC CAA ACA TGC	1154
Pro Gln Lys Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys	
345 350 355	
AGC TGT TAC AGA CGG CCA TGT ACG AAC CGC CAG AAG GCT TGT GAG CCA	1202
Ser Cys Tyr Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro	
360 365 370	
GGA TTT TCA TAT AGT GAA GAA GTG TGT CGT TGT GTC CCT TCA TAT TGG	1250
Gly Phe Ser Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp	
375 380 385 390	
CAA AGA CCA CAA ATG AGC TAAGATTGTA CTGTTTTCCA GTTCATCGAT	1298
Gln Arg Pro Gln Met Ser	
395	
TTTCTATTAT GGAAAACTGT GTTGCCACAG TAGAACTGTC TGTGAACAGA GAGACCCTTG	1358
TGGGTCCATG CTAACAAAGA CAAAAGTCTG TCTTTCCTGA ACCATGTGGA TAACTTTACA	1418
GAAATGGACT GGAGCTCATC TGCAAAAGGC CTCTTGTAAG GACTGGTTTT CTGCCAATGA	1478
CCAAACAGCC AAGATTTTCC TCTTGTGATT TCTTTAAAAG AATGACTATA TAATTTATTT	1538
CCACTAAAAA TATTGTTTCT GCATTCATTT TTATAGCAAC AACAATTGGT AAAACTCACT	1598
GTGATCAATA TTTTATATC ATGCAAAATA TGTTTAAAAT AAAATGAAAA TTGTATTTAT	1658
AAAAAAAAAA AAAAAA	1674

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

-5-

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met His Ser Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala  
 -23                    -20                    -15                    -10

Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Phe  
                   -5                    1                    5

Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala  
 10                    15                    20                    25

Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser  
                   30                    35                    40

Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met  
                   45                    50                    55

Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln  
                   60                    65                    70

Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala  
                   75                    80                    85

His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys  
 90                    95                    100                    105

Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe  
                   110                    115                    120

Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr  
                   125                    130                    135

Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr  
                   140                    145                    150

Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu  
                   155                    160                    165

Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser  
 170                    175                    180                    185

Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile  
                   190                    195                    200

Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn  
                   205                    210                    215

Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys  
                   220                    225                    230

Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser  
                   235                    240                    245

Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu  
 250                    255                    260                    265

-6-

Glu	Thr	Cys	Gln	Cys	Val	Cys	Arg	Ala	Gly	Leu	Arg	Pro	Ala	Ser	Cys		
270								275						280			
Gly	Pro	His	Lys	Glu	Leu	Asp	Arg	Asn	Ser	Cys	Gln	Cys	Val	Cys	Lys		
285								290						295			
Asn	Lys	Leu	Phe	Pro	Ser	Gln	Cys	Gly	Ala	Asn	Arg	Glu	Phe	Asp	Glu		
300								305						310			
Asn	Thr	Cys	Gln	Cys	Val	Cys	Lys	Arg	Thr	Cys	Pro	Arg	Asn	Gln	Pro		
315								320						325			
Leu	Asn	Pro	Gly	Lys	Cys	Ala	Cys	Glu	Cys	Thr	Glu	Ser	Pro	Gln	Lys		
330								335						340		345	
Cys	Leu	Leu	Lys	Gly	Lys	Lys	Phe	His	His	Gln	Thr	Cys	Ser	Cys	Tyr		
350								355						360			
Arg	Arg	Pro	Cys	Thr	Asn	Arg	Gln	Lys	Ala	Cys	Glu	Pro	Gly	Phe	Ser		
365								370						375			
Tyr	Ser	Glu	Glu	Val	Cys	Arg	Cys	Val	Pro	Ser	Tyr	Trp	Gln	Arg	Pro		
380								385						390			
Gln	Met	Ser															
395																	

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1526 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- ```
(A) NAME/KEY: sig_peptide
(B) LOCATION: 71..142
```

(ix) FEATURE:

- (A) NAME/KEY: mat\_peptide  
(B) LOCATION: 143..1120

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 71..1120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGAGGCCACG GCTTATGCAA GCAAAGATCT GGAGGAGCAG TTACGGTCTG TGTCCAGTGT

-7-

|                                                                 |     |
|-----------------------------------------------------------------|-----|
| AGATGAACTC ATG ACT GTA CTC TAC CCA GAA TAT TGG AAA ATG TAC AAG  | 109 |
| Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Tyr Lys             |     |
| -24 -20 -15                                                     |     |
| TGT CAG CTA AGG AAA GGA GGC TGG CAA CAT AAC AGA GAA CAG GCC AAC | 157 |
| Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln Ala Asn |     |
| -10 -5 1 5                                                      |     |
| CTC AAC TCA AGG ACA GAA GAG ACT ATA AAA TTT GCT GCA GCA CAT TAT | 205 |
| Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr |     |
| 10 15 20                                                        |     |
| AAT ACA GAG ATC TTG AAA AGT ATT GAT AAT GAG TGG AGA AAG ACT CAA | 253 |
| Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln |     |
| 25 30 35                                                        |     |
| TGC ATG CCA CGG GAG GTG TGT ATA GAT GTG GGG AAG GAG TTT GGA GTC | 301 |
| Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Val |     |
| 40 45 50                                                        |     |
| GCG ACA AAC ACC TTC TTT AAA CCT CCA TGT GTG TCC GTC TAC AGA TGT | 349 |
| Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys |     |
| 55 60 65                                                        |     |
| GGG GGT TGC TGC AAT AGT GAG GGG CTG CAG TGC ATG AAC ACC AGC ACG | 397 |
| Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr Ser Thr |     |
| 70 75 80 85                                                     |     |
| AGC TAC CTC AGC AAG ACG TTA TTT GAA ATT ACA GTG CCT CTC TCT CAA | 445 |
| Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln |     |
| 90 95 100                                                       |     |
| GGC CCC AAA CCA GTA ACA ATC AGT TTT GCC AAT CAC ACT TCC TGC CGA | 493 |
| Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg |     |
| 105 110 115                                                     |     |
| TGC ATG TCT AAA CTG GAT GTT TAC AGA CAA GTT CAT TCC ATT ATT AGA | 541 |
| Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg |     |
| 120 125 130                                                     |     |
| CGT TCC CTG CCA GCA ACA CTA CCA CAG TGT CAG GCA GCG AAC AAG ACC | 589 |
| Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr |     |
| 135 140 145                                                     |     |
| TGC CCC ACC AAT TAC ATG TGG AAT AAT CAC ATC TGC AGA TGC CTG GCT | 637 |
| Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys Leu Ala |     |
| 150 155 160 165                                                 |     |
| CAG GAA GAT TTT ATG TTT TCC TCG GAT GCT GGA GAT GAC TCA ACA GAT | 685 |
| Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser Thr Asp |     |
| 170 175 180                                                     |     |
| GGG TTC CAT GAC ATC TGT GGA CCA AAC AAG GAG CTG GAT GAA GAG ACC | 733 |
| Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr |     |
| 185 190 195                                                     |     |

-8-

|                                                                                                                                                       |      |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| TGT CAG TGT GTC TGC AGA GCG GGG CTT CGG CCT GCC AGC TGT GGA CCC<br>Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys Gly Pro<br>200 205 210     | 781  |
| CAC AAA GAA CTA GAC AGA AAC TCA TGC CAG TGT GTC TGT AAA AAC AAA<br>His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys Asn Lys<br>215 220 225     | 829  |
| CTC TTC CCC AGC CAA TGT GGG GCC AAC CGA GAA TTT GAT GAA AAC ACA<br>Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr<br>230 235 240 245 | 877  |
| TGC CAG TGT GTA TGT AAA AGA ACC TGC CCC AGA AAT CAA CCC CTA AAT<br>Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn<br>250 255 260     | 925  |
| CCT GGA AAA TGT GCC TGT GAA TGT ACA GAA AGT CCA CAG AAA TGC TTG<br>Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys Cys Leu<br>265 270 275     | 973  |
| TTA AAA GGA AAG AAG TTC CAC CAC CAA ACA TGC AGC TGT TAC AGA CGG<br>Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg<br>280 285 290     | 1021 |
| CCA TGT ACG AAC CGC CAG AAG GCT TGT GAG CCA GGA TTT TCA TAT AGT<br>Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser Tyr Ser<br>295 300 305     | 1069 |
| GAA GAA GTG TGT CGT TGT GTC CCT TCA TAT TGG CAA AGA CCA CAA ATG<br>Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Gln Arg Pro Gln Met<br>310 315 320 325 | 1117 |
| AGC TAAGATTGTA CTGTTTTCCA GTTCATCGAT TTTCTATTAT GGAAAACTGT<br>Ser                                                                                     | 1170 |
| GTTGCCACAG TAGAACTGTC TGTGAACAGA GAGACCCTTG TGGGTCCATG CTAACAAAGA                                                                                     | 1230 |
| CAAAAGTCTG TCTTTCCTGA ACCATGTGGA TAACTTTACA GAAATGGACT GGAGCTCATC                                                                                     | 1290 |
| TGCAAAAGGC CTCTTGTAAG GACTGGTTTT CTGCCAATGA CCAAACAGCC AAGATTTTCC                                                                                     | 1350 |
| TCTTGTGATT TCTTTAAAG AATGACTATA TAATTTATTT CCACTAAAAA TATTGTTTCT                                                                                      | 1410 |
| GCATTCATTT TTATAGCAAC AACAATTGGT AAAACTCACT GTGATCAATA TTTTATATC                                                                                      | 1470 |
| ATGCAAAATA TGTTTAAAT AAAATGAAAA TTGTATTTAT AAAAAAAAAA AAAAAA                                                                                          | 1526 |

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear



-9-

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met Tyr Lys Cys Gln Leu
-24          -20          -15          -10

Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln Ala Asn Leu Asn Ser
      -5              1              5

Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala His Tyr Asn Thr Glu
  10              15              20

Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys Thr Gln Cys Met Pro
  25              30              35              40

Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe Gly Val Ala Thr Asn
      45              50              55

Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr Arg Cys Gly Gly Cys
      60              65              70

Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr Ser Thr Ser Tyr Leu
      75              80              85

Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu Ser Gln Gly Pro Lys
      90              95              100

Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser Cys Arg Cys Met Ser
  105              110              115              120

Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile Ile Arg Arg Ser Leu
      125              130              135

Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn Lys Thr Cys Pro Thr
      140              145              150

Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys Leu Ala Gln Glu Asp
  155              160              165

Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser Thr Asp Gly Phe His
  170              175              180

Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu Glu Thr Cys Gln Cys
  185              190              195              200

Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys Gly Pro His Lys Glu
      205              210              215

Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys Asn Lys Leu Phe Pro
      220              225              230

Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu Asn Thr Cys Gln Cys
  235              240              245

```

-10-

Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro Leu Asn Pro Gly Lys  
 250 255 260

Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys Cys Leu Leu Lys Gly  
 265 270 275 280

Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr Arg Arg Pro Cys Thr  
 285 290 295

Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser Tyr Ser Glu Glu Val  
 300 305 310

Cys Arg Cys Val Pro Ser Tyr Trp Gln Arg Pro Gln Met Ser  
 315 320 325

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 196 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Arg Thr Leu Ala Cys Leu Leu Leu Leu Gly Cys Gly Tyr Leu Ala  
 1 5 10 15

His Val Leu Ala Glu Glu Ala Glu Ile Pro Arg Glu Val Ile Glu Arg  
 20 25 30

Leu Ala Arg Ser Gln Ile His Ser Ile Arg Asp Leu Gln Arg Leu Leu  
 35 40 45

Glu Ile Asp Ser Val Gly Ser Glu Asp Ser Leu Asp Thr Ser Leu Arg  
 50 55 60

Ala His Gly Val His Ala Thr Lys His Val Pro Glu Lys Arg Pro Leu  
 65 70 75 80

Pro Ile Arg Arg Lys Arg Ser Ile Glu Glu Ala Val Pro Ala Val Cys  
 85 90 95

Lys Thr Arg Thr Val Ile Tyr Glu Ile Pro Arg Ser Gln Val Asp Pro  
 100 105 110

Thr Ser Ala Asn Phe Leu Ile Trp Pro Pro Cys Val Glu Val Lys Arg  
 115 120 125

Cys Thr Gly Cys Cys Asn Thr Ser Ser Val Lys Cys Gln Pro Ser Arg  
 130 135 140

-11-

Val His His Arg Ser Val Lys Val Ala Lys Val Glu Tyr Val Arg Lys  
 145 150 155 160

Lys Pro Lys Leu Lys Glu Val Gln Val Arg Leu Glu Glu His Leu Glu  
 165 170 175

Cys Ala Cys Ala Thr Thr Ser Leu Asn Pro Asp Tyr Arg Glu Glu Asp  
 180 185 190

Thr Asp Val Arg  
 195

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 241 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Arg Cys Trp Ala Leu Phe Leu Ser Leu Cys Cys Tyr Leu Arg  
 1 5 10 15

Leu Val Ser Ala Glu Gly Asp Pro Ile Pro Glu Glu Leu Tyr Glu Met  
 20 25 30

Leu Ser Asp His Ser Ile Arg Ser Phe Asp Asp Leu Gln Arg Leu Leu  
 35 40 45

His Gly Asp Pro Gly Glu Glu Asp Gly Ala Glu Leu Asp Leu Asn Met  
 50 55 60

Thr Arg Ser His Ser Gly Gly Glu Leu Glu Ser Leu Ala Arg Gly Arg  
 65 70 75 80

Arg Ser Leu Gly Ser Leu Thr Ile Ala Glu Pro Ala Met Ile Ala Glu  
 85 90 95

Cys Lys Thr Arg Thr Glu Val Phe Glu Ile Ser Arg Arg Leu Ile Asp  
 100 105 110

Arg Thr Asn Ala Asn Phe Leu Val Trp Pro Pro Cys Val Glu Val Gln  
 115 120 125

Arg Cys Ser Gly Cys Cys Asn Asn Arg Asn Val Gln Cys Arg Pro Thr  
 130 135 140

Gln Val Gln Leu Arg Pro Val Gln Val Arg Lys Ile Glu Ile Val Arg  
 145 150 155 160

[illegible]

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 232 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: not relevant

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Phe | Leu | Leu | Ser | Trp | Val | His | Trp | Ser | Leu | Ala | Leu | Leu | Leu |
| 1   |     |     | 5   |     |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Tyr | Leu | His | His | Ala | Lys | Trp | Ser | Gln | Ala | Ala | Pro | Met | Ala | Glu | Gly |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Gly | Gly | Gln | Asn | His | His | Glu | Val | Val | Lys | Phe | Met | Asp | Val | Tyr | Gln |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Arg | Ser | Tyr | Cys | His | Pro | Ile | Glu | Thr | Leu | Val | Asp | Ile | Phe | Gln | Glu |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Tyr | Pro | Asp | Glu | Ile | Glu | Tyr | Ile | Phe | Lys | Pro | Ser | Cys | Val | Pro | Leu |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Met | Arg | Cys | Gly | Gly | Cys | Cys | Asn | Asp | Glu | Gly | Leu | Glu | Cys | Val | Pro |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Thr | Glu | Glu | Ser | Asn | Ile | Thr | Met | Gln | Ile | Met | Arg | Ile | Lys | Pro | His |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Gln | Gly | Gln | His | Ile | Gly | Glu | Met | Ser | Phe | Leu | Gln | His | Asn | Lys | Cys |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

-13-

Glu Cys Arg Pro Lys Lys Asp Arg Ala Arg Gln Glu Lys Lys Ser Val  
 130 135 140  
 Arg Gly Lys Gly Lys Gly Gln Lys Arg Lys Arg Lys Lys Ser Arg Tyr  
 145 150 155 160  
 Lys Ser Trp Ser Val Tyr Val Gly Ala Arg Cys Cys Leu Met Pro Trp  
 165 170 175  
 Ser Leu Pro Gly Pro His Pro Cys Gly Pro Cys Ser Glu Arg Arg Lys  
 180 185 190  
 His Leu Phe Val Gln Asp Pro Gln Thr Cys Lys Cys Ser Cys Lys Asn  
 195 200 205  
 Thr Asp Ser Arg Cys Lys Ala Arg Gln Leu Glu Leu Asn Glu Arg Thr  
 210 215 220  
 Cys Arg Cys Asp Lys Pro Arg Arg  
 225 230

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 14 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Pro Xaa Cys Val Xaa Xaa Xaa Arg Cys Xaa Gly Cys Cys Asn  
 1 5 10

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGCTTCCGG CTCGTATG

18

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:

-14-

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGTTTTCCC AGTCACGAC

18

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCACATGGTT CAGGAAAGAC A

21

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TGTAATACGA CTCACTATAG GGATCCCGCC ATGGAGGCCA CGGCTTATGC

50

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

-15-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GATCTCTAGA TTAGCTCATT TGTGGTCT

28

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCGGATCCA TGACTGTACT CTACCCA

27

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGCTCTAGAT CAAGCGTAGT CTGGGACGTC GTATGGGTAC TCGAGGCTCA TTTGTGGTCT

60

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3974 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGTACCTAAG TGAGTAGGGC GTCCGATCGA CGGACGCCTT TTTTTTGAAT TCGTAATCAT

60

GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG

120

CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG

180

CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA

240

-16-

|                                                                   |      |
|-------------------------------------------------------------------|------|
| TCGGCCAACG CGCGGGGAGA GCGGTTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA | 300  |
| CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG | 360  |
| TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC | 420  |
| AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC | 480  |
| CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC | 540  |
| TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC | 600  |
| TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA | 660  |
| GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTTC CTCCAAGCTG GGCTGTGTGC | 720  |
| ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA | 780  |
| ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG | 840  |
| CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA | 900  |
| GAAGAACAGT ATTTGGTATC TCGCTCTGCG TGAAGCCAGT TACCTTCGGA AAAAGAGTTG | 960  |
| GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC | 1020 |
| AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT | 1080 |
| CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCGTCGA | 1140 |
| CAATTCGCGC GCGAAGGCGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAAA | 1200 |
| ACCTTTCGCG GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAGGGT GGTGAATGTG | 1260 |
| AAACCAGTAA CGTTATACGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTTCC | 1320 |
| CGCGTGGTGA ACCAGGCCAG CCACGTTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGGCG | 1380 |
| ATGGCGGAGC TGAATTACAT TCCCAACCGC GTGGCACAAC AACTGGCGGG CAAACAGTCG | 1440 |
| TTGCTGATTG GCGTTGCCAC CTCCAGTCTG GCCCTGCAGC CGCCGTCGCA AATTGTCGCG | 1500 |
| GCGATTAAAT CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTCGAT GGTAGAACGA | 1560 |
| AGCGGCGTCG AAGCCTGTAA AGCGGCGGTG CACAATCTTC TCGCGCAACG CGTCAGTGGG | 1620 |
| CTGATCATTG ACTATCCGCT GGATGACCAG GATGCCATTG CTGTGGAAGC TGCCTGCACT | 1680 |
| AATGTTCCGG CGTTATTTCT TGATGTCTCT GACCAGACAC CCATCAACAG TATTATTTTC | 1740 |
| TCCCATGAAG ACGGTACGCG ACTGGGCGTG GAGCATCTGG TCGCATTGGG TCACCAGCAA | 1800 |
| ATCGCGCTGT TAGCGGGCCC ATTAAGTTCT GTCTCGGCGC GTCTGCGTCT GGCTGGCTGG | 1860 |
| CATAAATATC TCACTCGCAA TCAAAATCAG CCGATAGCGG AACGGGAAGG CGACTGGAGT | 1920 |



-17-

|            |            |             |             |            |            |      |
|------------|------------|-------------|-------------|------------|------------|------|
| GCCATGTCCG | GTTTTCAACA | AACCATGCAA  | ATGCTGAATG  | AGGGCATCGT | TCCCCTGCG  | 1980 |
| ATGCTGGTTG | CCAACGATCA | GATGGCGCTG  | GGCGCAATGC  | GCGCCATTAC | CGAGTCCGGG | 2040 |
| CTGCGCGTTG | GTGCGGATAT | CTCGGTAGTG  | GGATACGACG  | ATACCGAAGA | CAGCTCATGT | 2100 |
| TATATCCCGC | CGTTAACCAC | CATCAAACAG  | GATTTTCGCC  | TGCTGGGGCA | AACCAGCGTG | 2160 |
| GACCGCTTGC | TGCAACTCTC | TCAGGGCCAG  | GCGGTGAAGG  | GCAATCAGCT | GTTGCCCGTC | 2220 |
| TCACTGGTGA | AAAGAAAAAC | CACCCCTGGCG | CCCAATACGC  | AAACCGCCTC | TCCCCGCGCG | 2280 |
| TTGGCCGATT | CATTAATGCA | GCTGGCACGA  | CAGGTTTCCC  | GACTGGAAAG | CGGGCAGTGA | 2340 |
| GCGCAACGCA | ATTAATGTAA | GTTAGCGCGA  | ATTGTCGACC  | AAAGCGGCCA | TCGTGCCTCC | 2400 |
| CCACTCCTGC | AGTTCGGGGG | CATGGATGCG  | CGGATAGCCG  | CTGCTGGTTT | CCTGGATGCC | 2460 |
| GACGGATTTG | CACTGCCGGT | AGAACTCCGC  | GAGGTCGTCC  | AGCCTCAGGC | AGCAGCTGAA | 2520 |
| CCAACTCGCG | AGGGGATCGA | GCCCCGGGGT  | GGCGAAGAAC  | TCCAGCATGA | GATCCCCGCG | 2580 |
| CTGGAGGATC | ATCCAGCCGG | CGTCCCGGAA  | AACGATTCCG  | AAGCCCAACC | TTTCATAGAA | 2640 |
| GGCGGCGGTG | GAATCGAAAT | CTCGTGATGG  | CAGGTTGGGC  | GTCGCTTGGT | CGGTCAATTC | 2700 |
| GAACCCCAAG | GTCCCGCTCA | GAAGAACTCG  | TCAAGAAGGC  | GATAGAAGGC | GATGCGCTGC | 2760 |
| GAATCGGGAG | CGGCGATACC | GTAAAGCACG  | AGGAAGCGGT  | CAGCCCATTC | GCCGCCAAGC | 2820 |
| TCTTCAGCAA | TATCACGGGT | AGCCAACGCT  | ATGTCCTGAT  | AGCGGTCCGC | CACACCCAGC | 2880 |
| CGGCCACAGT | CGATGAATCC | AGAAAAGCGG  | CCATTTTCCA  | CCATGATATT | CGGCAAGCAG | 2940 |
| GCATCGCCAT | GGGTCACGAC | GAGATCCTCG  | CCGTCGGGCA  | TGCGCGCCTT | GAGCCTGGCG | 3000 |
| AACAGTTCGG | CTGGCGCGAG | CCCCTGATGC  | TCTTCGTCCA  | GATCATCCTG | ATCGACAAGA | 3060 |
| CCGGCTTCCA | TCCGAGTACG | TGCTCGCTCG  | ATGCGATGTT  | TCGCTTGGTG | GTCGAATGGG | 3120 |
| CAGGTAGCCG | GATCAAGCGT | ATGCAGCCGC  | CGCATTG CAT | CAGCCATGAT | GGATACTTTC | 3180 |
| TCGGCAGGAG | CAAGGTGAGA | TGACAGGAGA  | TCCTGCCCCG  | GCACTTCGCC | CAATAGCAGC | 3240 |
| CAGTCCCTTC | CCGCTTCAGT | GACAACGTCG  | AGCACAGCTG  | CGCAAGGAAC | GCCCGTCGTG | 3300 |
| GCCAGCCACG | ATAGCCGCGC | TGCCTCGTCC  | TGCAGTTCAT  | TCAGGGCACC | GGACAGGTCG | 3360 |
| GTCTTGACAA | AAAGAACCGG | GCGCCCCCTG  | GCTGACAGCC  | GGAACACGGC | GGCATCAGAG | 3420 |
| CAGCCGATTG | TCTGTTGTGC | CCAGTCATAG  | CCGAATAGCC  | TCTCCACCCA | AGCGGCCGGA | 3480 |
| GAACCTGCGT | GCAATCCATC | TTGTTCAATC  | ATGCGAAACG  | ATCCTCATCC | TGTCTCTTGA | 3540 |
| TCAGATCTTG | ATCCCCTGCG | CCATCAGATC  | CTTGGCGGCA  | AGAAAGCCAT | CCAGTTTACT | 3600 |

-18-

```

TTGCAGGGCT TCCCAACCTT ACCAGAGGGC GCCCCAGCTG GCAATTCCGG TTCGCTTGCT      3660
GTCCATAAAA CCGCCCAGTC TAGCTATCGC CATGTAAGCC CACTGCAAGC TACCTGCTTT      3720
CTCTTTGCGC TTGCGTTTTC CCTTGTCAG ATAGCCCAGT AGCTGACATT CATCCGGGGT      3780
CAGCACCGTT TCTGCGGACT GGCTTTCTAC GTGTTCCGCT TCCTTTAGCA GCCCTTGCGC      3840
CCTGAGTGCT TCGGCAGCG TGAAGCTTAA AAAACTGCAA AAAATAGTTT GACTTGTGAG      3900
CGGATAACAA TTAAGATGTA CCCAATTGTG AGCGGATAAC AATTTCACAC ATTAAAGAGG      3960
AGAAATTACA TATG  3974

```

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 112 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: both

## (ii) MOLECULE TYPE: DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

AAGCTTAAAA AACTGCAAAA AATAGTTTGA CTTGTGAGCG GATAACAATT AAGATGTACC      60
CAATTGTGAG CGGATAACAA TTTCACACAT TAAAGAGGAG AAATTACATA TG              112

```

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 419 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

Met His Ser Leu Gly Phe Phe Ser Val Ala Cys Ser Leu Leu Ala Ala
1           5           10           15
Ala Leu Leu Pro Gly Pro Arg Glu Ala Pro Ala Ala Ala Ala Phe
20           25           30
Glu Ser Gly Leu Asp Leu Ser Asp Ala Glu Pro Asp Ala Gly Glu Ala
35           40           45

```

-19-

Thr Ala Tyr Ala Ser Lys Asp Leu Glu Glu Gln Leu Arg Ser Val Ser  
 50 55 60  
 Ser Val Asp Glu Leu Met Thr Val Leu Tyr Pro Glu Tyr Trp Lys Met  
 65 70 75 80  
 Tyr Lys Cys Gln Leu Arg Lys Gly Gly Trp Gln His Asn Arg Glu Gln  
 85 90 95  
 Ala Asn Leu Asn Ser Arg Thr Glu Glu Thr Ile Lys Phe Ala Ala Ala  
 100 105 110  
 His Tyr Asn Thr Glu Ile Leu Lys Ser Ile Asp Asn Glu Trp Arg Lys  
 115 120 125  
 Thr Gln Cys Met Pro Arg Glu Val Cys Ile Asp Val Gly Lys Glu Phe  
 130 135 140  
 Gly Val Ala Thr Asn Thr Phe Phe Lys Pro Pro Cys Val Ser Val Tyr  
 145 150 155 160  
 Arg Cys Gly Gly Cys Cys Asn Ser Glu Gly Leu Gln Cys Met Asn Thr  
 165 170 175  
 Ser Thr Ser Tyr Leu Ser Lys Thr Leu Phe Glu Ile Thr Val Pro Leu  
 180 185 190  
 Ser Gln Gly Pro Lys Pro Val Thr Ile Ser Phe Ala Asn His Thr Ser  
 195 200 205  
 Cys Arg Cys Met Ser Lys Leu Asp Val Tyr Arg Gln Val His Ser Ile  
 210 215 220  
 Ile Arg Arg Ser Leu Pro Ala Thr Leu Pro Gln Cys Gln Ala Ala Asn  
 225 230 235 240  
 Lys Thr Cys Pro Thr Asn Tyr Met Trp Asn Asn His Ile Cys Arg Cys  
 245 250 255  
 Leu Ala Gln Glu Asp Phe Met Phe Ser Ser Asp Ala Gly Asp Asp Ser  
 260 265 270  
 Thr Asp Gly Phe His Asp Ile Cys Gly Pro Asn Lys Glu Leu Asp Glu  
 275 280 285  
 Glu Thr Cys Gln Cys Val Cys Arg Ala Gly Leu Arg Pro Ala Ser Cys  
 290 295 300  
 Gly Pro His Lys Glu Leu Asp Arg Asn Ser Cys Gln Cys Val Cys Lys  
 305 310 315 320  
 Asn Lys Leu Phe Pro Ser Gln Cys Gly Ala Asn Arg Glu Phe Asp Glu  
 325 330 335  
 Asn Thr Cys Gln Cys Val Cys Lys Arg Thr Cys Pro Arg Asn Gln Pro

-20-

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 340                                                             | 345 | 350 |
| Leu Asn Pro Gly Lys Cys Ala Cys Glu Cys Thr Glu Ser Pro Gln Lys |     |     |
| 355                                                             | 360 | 365 |
| Cys Leu Leu Lys Gly Lys Lys Phe His His Gln Thr Cys Ser Cys Tyr |     |     |
| 370                                                             | 375 | 380 |
| Arg Arg Pro Cys Thr Asn Arg Gln Lys Ala Cys Glu Pro Gly Phe Ser |     |     |
| 385                                                             | 390 | 395 |
| Tyr Ser Glu Glu Val Cys Arg Cys Val Pro Ser Tyr Trp Gln Arg Pro |     |     |
| 405                                                             | 410 | 415 |
| Gln Met Ser                                                     |     |     |

## (2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GCAGCACATA TGACAGAAGA GACTATAAAA

30

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCAGCAGGTA CCTCACAGTT TAGACATGCA

30

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

-21-

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCAGCAGGTA CCTCAACGTC TAATAATGGA

30

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCAGCAGGAT CCCACAGAAG AGACTATAAA

30

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCAGCATCTA GATCACAGTT TAGACATGCA

30

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCAGCAGGAT CCCACAGAAG AGACTATAAA ATTTGCTGC

39

(2) INFORMATION FOR SEQ ID NO:25:

-22-

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATCTA GATCAACGTC TAATAATGGA ATGAAC

36

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 55 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GATCGATCCA TCATGCACTC GCTGGGCTTC TTCTCTGTGG CGTGTTCTCT GCTCG

55

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCAGGGTACG GATCCTAGAT TAGCTCATTT GTGGTCTTT

39

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

-23-

GACTGGATCC GCCACCATGC ACTCGCTGGG CTTCTTCTC

39

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GACTGGTACC TTATCACATA AAATCTTCCT GAGCC

35

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GACTGGATCC GCCACCATGC ACTCGCTGGG CTTCTTCTC

39

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GACTGGTACC TTATCAGTCT AGTTCTTTGT GGGG

34

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

-24-

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GACTGGATCC GCCACCATGC ACTCGCTGGG CTTCTTCTC

39

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GACTGGTACC TCATTACTGT GGACTTTCTG TACATTC

37

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GCAGCAGGAT CCACAGAAGA GACTATAAAA TTTGCTGC

38

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CGTCGTTCTA GATCACAGTT TAGACATGCA TCGGCAG

37